The Patent Office

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

CERTIFIED COPY OF
PRIORITY DOCUMENT

Signed

Dated 29 February 2000

Patents Form 1/77

Patents Act 1977
(Rule 16)

The Patent Office

26MAY98 E362875-1 C38944
P01/7700 25.00 - 9811130.5

THE PATENT OFFICE
K
23 MAY 1998
RECEIVED BY POST

23 MAY 1998

The Patent Office

Cardiff Road,
Newport
Gwent NP9 1RH

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

1. Your reference

2. Patent application number
(The Patent Office will fill in this part) 9811130.5

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

BIOVATION LIMITED
AURIS BUSINESS CENTRE
23 ST NINIAN DRIVE
ABERDEEN AB24 3RY
INCORP. IN SCOTLAND GB 4200903

Patents ADP number *(if you know it)*

If the applicant is a corporate body, give the country/state of its incorporation

4. Title of the invention

DISEASE ASSOCIATED PROTEIN SCREENING ~~MODIFICATIONS~~

5. Name of your agent *(if you have one)*

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

~~AGENTS NAME~~
Kilburn & Strode
20 Red Lion Street
London
WC1R 4PJ

See 61/77 filed 22/9/98.

Patents ADP number *(if you know it)* 125001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*
—		

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*
—	

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*

a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

YES

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	
Description	10
Claim(s)	—
Abstract	—
Drawing(s)	—

10. If you are also filing any of the following, state how many against each item.

Priority documents	—
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	

11. I/We request the grant of a patent on the basis of this application.

Signature

Date 22/5/98

12. Name and daytime telephone number of person to contact in the United Kingdom

01224 273406

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.

b) Write your answers in capital letters using black ink or you may type them.

c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) For details of the fee and ways to pay please contact the Patent Office.

Disease Associated Protein Screening

The present invention relates to new screening methods to detect polypeptides and to detect polypeptide modifications associated with disease or drug treatments or associated with variation between different individuals. In particular, the invention relates to methods whereby libraries of genes are used as the basis for forming arrays of proteins in order to screen for detect polypeptides and to detect polypeptide modifications genes. The invention also relates to specific screening methods for novel polypeptide interactions especially protein-protein binding.

From the ongoing initiative to sequence the human genome has developed a major initiative to find polymorphisms or mutations in human genes which might have causal relationships with human diseases and, in addition, to alterations in the expression of genes which might relate to disease. This has given rise to many high-throughput screening technologies in order to find such polymorphisms, mutations or alterations in gene expression. In addition, such analyses are now being undertaken to analyse human responses to certain treatments and therefore to predict responsiveness to these treatments. Commonly, polymorphisms, mutations and gene expression changes are reflected in proteins produced by these genes and it is the actions of these proteins which can directly determine biological outcomes such as development of a disease or response to a treatment. In addition, most eukaryotic proteins are modified post-translationally and such modifications cannot be analysed at the gene level and it therefore follows that analysis of the screening for proteins and protein modifications is more likely to give a precise relationship of changes with disease compared to analysis of DNA or RNA is likely The analysis of protein associations in disease has, to date, been dominated by 2D (two-dimensional) protein gel analysis from human tissue or cells. For this technology, cellular proteins are usually separated on the basis of charge in one dimension and on the basis of size in the other dimension. Proteins can either be identified with reference to the electophoresis migration pattern of a known protein or by elution of the protein from the electrophoretically separated spot and analysis by methods such as mass spectrometry and nuclear magnetic resonance. However, limitations of the 2D protein gel method include the limited resolution and detection of proteins from a cell (typically only 5000 cellular proteins are clearly detected), the limitation to identification of separated proteins (for example, mass spectrometry usually requires 100fmoles or more of protein for identification), and the specialist nature of the technique. In addition, the analysis of post-translational modification of proteins by 2D gel analysis is limited and protein-protein binding interactions cannot be detected by this method.

A large number of cellular processes are controlled by the transient interaction between a modifying enzyme and its protein substrate. Such interactions are difficult to detect by conventional methods as the proteins and conditions required for the reaction have first to be identified. In addition the interaction often relies on native protein conformation which is often not achieved during *in vitro* translation. Protein modifying enzymes such as kinases, phosphatases, transferases, proteases etc. control all manner of fundamental cellular processes (Pawson, *Nature*, **373** (1995), pp573) and have also been shown to be involved in disease pathways.

These transient protein/protein interactions can result in a number of different effects some of which can be measured. The kinetic properties of a protein can be altered resulting in altered binding of substrates (Prelich *et al.*, *Nature* **326**, (1989) pp 517) or altered catalysis (Porpaczy *et al.*, *Biochim. Biophys. Acta*, **749** (1983), pp 172). Protein/protein interactions can cause the formation of a new binding site e.g. an ATP binding site is formed by the interaction of the α and β subunits of the *E. coli* ATPase (Weber *et al.*, *J. Biol. Chem.*, **268**, (1993), pp 6241). Substrate specificity of a protein can be altered by protein/protein interactions as exemplified by the interaction of different transcription factors with RNA polymerase directing the polymerase to specific promoters (White & Jackson, *Trends Genet.*, **8**, (1988), pp 284). Alternatively protein/protein interactions can cause inactivation, for example when a protein interacts with an inhibitor (Vincent & Lazdunski, *Biochemistry*, **11**, (1972), pp2967).

For the purpose of the present invention, "synthetic proteins" produced by *in vitro* methods such as *in vitro* transcription and translation, ribosome display and phage display shall be distinguished from "natural proteins" derived directly from tissue or cell extracts.

The present invention provides novel methods for screening for proteins and post-translational protein modifications using gene libraries as the starting point for synthesis of synthetic proteins. In the principal aspect of the invention, these gene libraries provide the basis for segregating individual proteins or groups of proteins in order to detect polypeptides and to detect polypeptide modifications. The method principally avoids the use of protein gel electrophoresis. The invention provides methods for the high-throughput analysis of proteins in normal and disease tissues or cells principally to detect differences between normal and disease proteins and also for the analysis of protein changes in relation to drug and other disease treatments. Furthermore, the method also provides for the protein "fingerprinting" of individuals as an alternative to genetic fingerprinting in order to identify individuals and different tissues or cells. It is an important aspect of the present invention that proteins are analysed either in pools or in arrays of individual proteins whereby, once modifications are detected using a variety of methods, the identity of the protein can be determined either directly or indirectly by either its location in a particular pool or its location at a particular position in an array. Equally, it is an important aspect of the present invention that the methods allow for the comparative detection of differences in proteins or protein modifications between different tissues or cells and therefore the methods constitute a screen for these differences.

The method is based primarily on the use of synthetic proteins produced by recombinant DNA methods, especially proteins produced from one or more genes by *in vitro* transcription and translation (IVTT). A particular aspect of the present invention is the use of "polypeptide display" methods for isolating polypeptides with useful phenotypes from gene libraries encoding a large mixture of polypeptides whereby the corresponding genes can be recovered swiftly and easily. *In vitro* methods of polypeptide display have been developed which allow greater scope for the type of protein displayed than cell based display systems and in particular display of proteins associated with ribosomes ("ribosome display") which can provide proteins linked via the ribosome complex to corresponding mRNAs which can then be

sequenced in order to reveal the identity of the protein. The present method can also incorporate other protein display methods such as phage display.

In one embodiment of the present method, protein modifications are screened by creation of a gene library which is then use to create synthetic proteins for example by IVTT to successively produce mRNA and protein from such vectors preferably under such conditions as to maximise yield of full-length synthetic proteins which have disassociated from mRNA. Through the incorporation of modified amino acids such as biotinyl-lysine or the incorporation of an amino acid sequence tag encoded by the plasmid vector or introduced by PCR, or through chemical modification of the protein, the synthetic protein can next be immobilised onto a solid phase such as a magnetic bead via a biotin binding protein (e.g. streptavidin) or via a sequence tag binding protein such as an anti-tag antibody. In a screen for protein modification, through appropriate distribution of the starting genes, small pools of proteins or individual proteins are produced in arrays such as in microtitre wells and immobilised in these arrays. Alternatively, the synthetic proteins might be produced in arrays as individual small pools of proteins or individual proteins and then transferred individually (either manually or by use of an automatic dispenser) to different loci on a flat solid phase such as on a glass "chip" whereby the proteins interact with functional moieties on the solid phase resulting in immobilisation. Next, these synthetic proteins are treated with extracts of tissues or cells whereby enzymes present in these tissues or cells can effect modifications to the synthetic proteins. Finally, such modifications are then be detected using specific detection systems such as fluorescently labelled anti-phoshotyrosine antibodies. If the signal from the detection system is compared from replicate arrays of synthetic proteins which are modified by different extracts of tissues or cells, then differences in these signals between replicate synthetic proteins will indicate differences in the protein modification activities from these different tissues or cells.

In another embodiment of the present method, protein-protein binding interactions are screened by creation of a gene library which is then used to create synthetic proteins such as by IVTT, preferably in such vectors and under such conditions as to maximise yield of full-length synthetic primary proteins which have disassociated from mRNA. Through the incorporation of modified amino acids or an amino acid sequence tag, or through chemical modification of the protein, or by use of a chemically reactive solid phase, the synthetic primary protein can be immobilised onto a solid phase either immediately after synthesis or following the interaction with secondary proteins to form any protein-protein interactions. Through appropriate distribution of the starting genes, as in screens for protein modification, small pools of primary proteins or individual proteins are produced in arrays such as in microtitre wells or on glass chips and immobilised in these arrays. Next, these synthetic primary proteins can optionally be treated with extracts of tissues or cells whereby enzymes present in these tissues or cells can effect modifications to the synthetic proteins. Next, one or more (including a library of) secondary proteins are then added in order to form any suitable binding interactions of appropriate secondary proteins with the primary proteins. Optionally, these secondary proteins may be synthetic and displayed on ribosomes or phage such that the identity of the secondary protein can subsequently be determined. Alternatively, the secondary proteins could be natural proteins derived from tissues or cells. Protein-protein binding interactions between primary and secondary proteins can

then be detected by a range of methods. For example, the primary protein can be immobilised onto a solid phase and the presence of a synthetic secondary protein determined via an amino acid sequence tag incorporated into the secondary protein with detection using specific detection systems such as fluorescently labelled anti-tag antibody. Alternatively, if the secondary protein is associated with its coding nucleic acid such as for Ribosome or phage display, then the secondary protein can be detected by a PCR assay for the associated nucleic acid. Alternatively, binding of the secondary protein may be detected through the action of the secondary protein in blocking a site on the synthetic primary protein which is detected in the assay system and which becomes blocked if a secondary protein binds. Alternatively, if the secondary protein is a natural protein, it may be labelled directly after binding, for example by addition of a UV-activatable biotin molecule, and detected directly, or it may be laser vapourised and detected by MALDI (Matrix Assisted Laser Desorption Ionisation) and mass spectrometry. If the binding of secondary proteins to replicate arrays of synthetic primary proteins is compared between different extracts of tissues or cells, then differences in these signals between replicate synthetic proteins will indicate differences in the protein-protein binding activities from these different tissues or cells.

It will be obvious that by detecting a range of protein modifications and by detecting protein-protein binding with a common detection format such as by using fluorescent labelled antibodies to protein tags or epitopes, then simultaneous parallel screening for a number of different protein modifications could be undertaken in order to screen for differences between different tissues and cells.

In another embodiment of the present invention, protein-protein interactions are used as the basis for forming arrays of secondary proteins, especially natural tissue or cell derived proteins which normally would be very difficult to comprehensively array using standard technology. The primary proteins comprise a highly diverse set of proteins such as immunoglobulins which provide a diverse set of binding specificities for a set of secondary proteins such as natural proteins from tissue or cell extracts or synthetic proteins from a ribosome or phage displayed protein library. As an example, the primary proteins would derive from a Ribosome displayed single-chain antibody (SCA) variable region library formed by direct PCR amplification of immunoglobulin variable regions from mammalian B lymphocytes and cloning into suitable vectors for *in vitro* transcription. Members of the SCA library are arrayed either individually or in groups into microtitre wells or directly on a solid phase such as a glass chip and the SCA proteins are produced by IVTT and immobilised via a protein tag or by direct chemical coupling to a solid phase. Then a mixture of natural proteins from a tissue or cell extract or a mixture of synthetic proteins originating from a gene library (for example ribosome or phage displayed proteins) is added in bulk to the array of primary proteins such as SCAs whereby the binding specificity imparted by the primary proteins for the secondary proteins leads to arraying of the secondary proteins with either single secondary proteins or groups of secondary proteins bindign to individual SCAs. The achieved array of secondary proteins is compared between different extracts of tissues or cells, then differences in these signals between replicate synthetic proteins will indicate differences in the presence of proteins from these different tissues or cells. Alternatively, these proteins can also be tested for various modifications such a phosphorylation by using, for example, fluorescently labelled antibodies to detect these

modifications.

It is a particular aspect of the present invention that proteins, either synthetic or natural, are distributed into arrays of single or multiple proteins in order to provide for comparisons of the presence or modification of these individual single or multiple protein members of the array between different tissue or cell extracts. Whilst arrays can readily be performed in microtitre plates by distributing individual members of groups of clones from a cDNA library in the individual wells of the plate, this manual (or robot-directed) arraying technology results in the physical barrier of the walls of the wells between different samples and this makes it difficult to add subsequent solutions such as wash solutions and libraries of secondary proteins quickly and accurately. In addition, the use of microtitre plates provides a limit to the density of arrayed proteins, such limit being dictated by the density of wells in the microtire plates available. It is therefore desirable to have methods which do not require arraying of proteins without physical barriers between the proteins.

In another embodiment of the invention where arrays of proteins are formed without physical barriers, the spatial distribution of proteins into arrays is achieved either by directly spatially immobilising individual synthetic proteins derived from a gene library by, for example, Ribosome display or by spatially immobilising nucleic acids, especially synthetic DNA, to which nucleic acids associated with a synthetic protein library are annealed in order to indirectly spatially immobilise members of the protein library. For direct spatial immobilisation, either synthetic proteins derived from a tissue or cell gene library are immobilised or synthetic protein probes such as SCAs are immobilised such that a mixture of secondary proteins can then be added whereby these would be spatially distributed according to their binding specificities for individual protein probes. Thus, the invention provides for new types of protein chips (cf Hutchens & Yip, *Rapid Comms. Mass Spec.* 7, (1993), pp576). Immobilisation of synthetic proteins on the chip is achieved by either covalent or non-covalent linkage. Non-covalent attachment of the proteins is facilitated by labelling the proteins for example with biotin using a commercially available reagent such as Sulfo-SBED (Pierce & Warriner, Chester, UK) which is then reacted with avidin. Covalent attachment of the proteins can also be achieved by activating the proteins with reactive species to facilitate cross-linking by any of the conventional ways, such as those described in O'Sullivan *et al.*, (Anal. Biochem. 100 (1979),108). Alternative methods for directly binding synthetic proteins would include binding of Ribosome displayed proteins by immobilising the associated RNA using, for example, an RNA binding protein (such as HIV tat protein) or by annealing the RNA to synthetic DNA molecules on the solid phase.

For indirect immobilisation of proteins via nucleic acids, the associated mRNA in a protein-ribosome-mRNA complex may be produced with a nucleotide sequence tag originally encoded by the plasmid vector encoding the mRNA. The nucleotide sequence tag may be variable or randomised in the plasmid vector preparation, for example by producing the plasmid vector using a synthetic region encoded by a random oligonucleotide mixture (with appropriate ends which anneal to the vector) to place random a sequence tag at, for example, the 5' end of the mRNA. Following production of the protein-ribosome-mRNA complexes from the gene library, the mRNA sequence tags may then be annealed to an array of synthetic oligonucleotides

individually positioned over a solid surface such as a glass slide (a "DNA biochip"). In this manner, individual proteins may finally be positioned at specific locations on the biochip. The spatially arrayed proteins may then be analysed for modification by tissue or cell extracts using, for example, fluorescent antibodies to screen for these modifications. In addition, the arrays may be used for identification of proteins which bind or modify the individual proteins on the biochip. The identity of the modified proteins or the proteins involved in protein-protein interactions may subsequently be determined from the known sequence of the immobilised oligonucleotide on the chip whereby this sequence is used to probe the plasmid library (for example, by PCR) in order to identify the specific gene associated with the complementary sequence tag within the plasmid library. The identity of the proteins may alternatively be achieved by PCR amplification of the mRNA sequence associated with the individual protein. This embodiment therefore relates to a protein array biochip with, in principle, individual proteins located at individual loci on the chip. Alternatively, the identity of the protein may be determined if the primary or secondary synthetic protein is produced with a random set of protein sequence tags which can be interrogated using, for example, a mixture of antibodies against those sequence tags. If, for example, the synthetic proteins are produced in such a manner that one terminus comprises either a random or semi-random stretch of amino acids each of which can be detected by an antibody from a library of antibodies with appropriate labels, then the binding of one or more specific antibodies can determine the identity of the synthetic protein. For example, if one terminus of the protein contains a random 8 amino acid sequence tag and a library of SCAs has been constructed and enriched for SCAs binding to an equivalent mixture of synthetic peptides, then individual or small groups of SCAs will bind to specific peptides and the peptides (and therefore synthetic proteins) can be identified by knowledge of the specificity of individual SCAs.

Encompassed by the present invention are also methods where mixtures of modified proteins or interacting proteins are preselected prior to analysis of differences between different tissues or cells. For example, primary synthetic proteins created by display methods such as ribosome display where individual mRNA molecules encoding a protein remains attached to the protein are subjected to treatment by tissue or cell extracts and then subjected to selection using antibodies against these modifications immobilised on a solid phase. Thus, modified proteins with associated nucleic acids are isolated and the profile of proteins isolated can be determined by several means including a simple analysis of PCR products, for example by agarose gel electrophoresis, and comparison of results from treatments with different cells and tissues. Similarly, for primary natural proteins which are then used to "capture" secondary synthetic proteins by protein-protein binding, interacting proteins could be isolated using, for example, an antibody against a tag on the secondary synthetic protein and a profile of the interacting proteins determined again by analysis of PCR products from the nucleic acids associated with the synthetic protein.

Encompassed by the present invention are also methods where synthetic proteins generated from gene libraries are subjected to protein gel electophoresis for the analysis of protein modification or protein-protein binding. It is well documented that post translational modifications such as proteolysis or phosphorylation change the electrophoretic mobility of a protein (Phizicky & Fields *Microbiol. Rev.*, **59**, (1995), pp94). One embodiment of the invention therefore provides a method whereby disease

associated protein modifications would be detected using 2D gel electrophoresis. For example, a cDNA library for subsequent IVTT or ribosome display would be constructed. To facilitate subsequent purification of the proteins, a 3' tag such as His or Flag may be incorporated into the transcription vector used. This will be achieved by standard molecular biology techniques that will be familiar to those skilled in the art. To facilitate subsequent detection of the proteins, they may be labelled during the translation process with labels such as ^{35}S or biotin. Following translation, the proteins could if required be purified to remove ribosomes and other factors for example by passage of the translation mixture over an affinity matrix containing ligands such as anti-Flag antibodies or nickel where the proteins have 3' flag or ployhistidine tags or using anti-ribosome antibodies to remove ribosomal components onto a solid phase or by simple ultracentrifugation of ribosomes. Protein from a clinical sample (either unlabelled or labelled, for example through reaction with a fluorescent moiety) would then be incubated with the purified *in vitro* translated library to effect either modification or protein interactions withy the synthetic protein. The total reaction would then be separated by 2D gel electrophoresis (Cash, J. Chromatography 698 (1995), p203) and the resultant gels analysed using appropriate computer software such as Phoretix-2D (Phoretix International, Newcastle upon Tyne, UK). A control reaction would be performed such that the synthetic protein library would be reacted with total protein from an alternative clinical sample, for example a normal sample if the first sample is a diseased sample. Disease associated modifications or protein/protein interactions would be identified as unique bands seen with the diseased proteins and not seen in the control proteins. In order to facilitate rapid identification of the gene associated with protein identified by the 2D gel analysis, the synthetic protein library could initially be sub-divided into a number of pools, each pool containing a proportion of the library. This would be achieved by dividing the initial mixture of genes. Each pool would then be screened as described above to identify a pool containing a protein of interest. The transcription mix for that pool would then be sub-divided again and the screening process repeated. This procedure would be repeated using progressively smaller pools of ribosome displayed proteins until a single clone was identified which encoded the protein of interest. Alternatively, following identification of a potential disease associated protein by 2D gel analysis, the protein would be purified from the gel according to standard protocols (Hager & Burgess, *Anal. Biochem.*, 109, (1980), pp76) and the purified protein panned with a scFv library to identify an antibody(s) which recognises the protein. The antibody(s) would then be used to screen a synthetic protein library constructed from the cDNA of the diseased sample to identify the protein with its associated gene tag. As an alternative highly innovative method to identify the protein, each recombinant protein may be produced with an amino acid sequence tag originally encoded by the plasmid vector and subsequently incorporated into the protein, for example by incorporating a leader sequence into the cloned cDNAs derived from the vector or from a mixture of synthetic oligonucleotides used to copy the cDNA for subsequent cloning into the vector. The nucleotide sequence tag may be variable or randomised in the plasmid vector preparation, for example by producing the plasmid vector using a synthetic region encoded by a random oligonucleotide mixture (with appropriate ends which anneal to the vector) to place random a sequence tag at, for example, the 5' end of the mRNA. Following production of the protein mixture by, for example, IVTT or ribsome display and separation of these proteins by 2D gel electrophoresis, individual

proteins could then be analysed by probing with antibodies specific for the various permutations of sequences.

It will be obvious for the present invention that a prime application will be in the analysis of proteins, protein modifications and protein-protein interactions which relate to human disease or to the healthcare, individual or drug treatment status of humans. The invention incorporates the use of "clinical samples", such term referring to samples which may be tissue, blood or other which could be expected to be affected by the particular disease. The diseased samples refers to a sample which can be demonstrated to be affected by a disease, healthcare status or drug treatment status. The normal sample refers to a sample which is not affected by the disease and represents normal healthcare status, but may vary between individuals and may be relevant to the subsequent outcome of drug treatment. For comparative purposes in searching for disease associated proteins or protein modifications/interactions, the normal and diseased samples would ideally be matched to reduce population induced heterogeneity. This may be achieved by obtaining normal and diseased samples from a single patient (e.g. cancerous breast tissue and unaffected breast tissue) or alternatively by pooling diseased and normal samples from a number of different patients.

The generation of cDNA libraries from both normal and clinical samples is performed by standard techniques involving the isolation of mRNA followed by reverse transcription to generate cDNA (Sambrook *et al.*, *Molecular Cloning : a Laboratory manual*, 2nd Ed., Cold Spring Harbor Press, 1989) or alternatively may be achieved using commercially available kits e.g. PolyATract System (Promega, Southampton, UK). The cloning of the cDNA's into a vector suitable for generation of synthetic proteins such as *in vitro* display by ribosome display and the conditions for *in vitro* display of the proteins would be as previously described (for example in Hanes and Pluckthun, *Proc. Natl. Acad. Sci.* 94 (1997), p4937). The isolation of total protein from the samples would be achieved by standard methods as detailed Bollag & Edelstein (Protein Methods, Wiley-Leiss, 1991).

The invention is illustrated by but not limited by the following examples.

Example 1

The starting point was a cDNA library constructed from human colon (Clontech, UK). The cDNAs were recloned into a pGEMT7/SP6 plasmid (Promega, Southampton, UK) providing an upstream promoter for T7 RNA polymerase using long synthetic oligonucleotides and PCR to provide an upstream bacterial ribosome binding site, a downstream spacer derived from the M13 phage gene III and a 3' transcriptional terminator region from the *E. coli* lpp terminator.

In vitro transcription of the pT7 plasmids was performed using a RiboMAX™ kit (Promega, Southampton, UK) according to the manufacturer's instructions. The resultant mRNA was purified according to the manufacturer's protocol.

In vitro translation was performed using *E coli* S30 Extract System with the inclusion of ^{35}S methionine (Promega, Southampton, UK) according to the manufacturer's instructions. The ribosomes were dissociated from the proteins by treatment with

EDTA as described by Hanes & Pluckthun (ibid) and the ribosomes removed by centrifugation (100,000 g for 30 min).

A crude protein extract was prepared from a colorectal carcinoma sample (ref). This was then incubated with the *in vitro* translated proteins above for 30 mins at 37°C in the presence of 1 mg/ml vanadyl ribonucleoside complexes and 10 units RNasin. As a control the *in-vitro* translated proteins were incubated with a crude protein preparation from normal colon tissue. The protein mixture was then solubilised according to the protocol of Cash *et al.*, (*Electrophoresis* 18 (1997), p2580). The soluble proteins were then analysed by 2D PAGE (Cash *et al.* ibid) and the proteins detected by autoradiography (Patton *et al., BioTechniques* 8 (1990), p518). Images of the gels were captured using a video camera connected to an image analysis system and then transferred to Phoretix-2D for further analysis. Analysis was performed using Phoretix-2D systems according to the manufacturer's instructions.

A total of 12 protein spots were identified in the in the colorectal cancer reacted proteins that were not present in the control.

Example 2

Recombinant purified rat ERK-2 containing a poly histidine tag fused to the amino terminus of the protein was purchased from Stratagene (cat # 20612). This product is the non-activated form of the protein and contains no phosphorylated residues. 0.5ug non-actiivated ERK-2 was reacted with a 5ul of cellular lysate (containing 2mg/ml total protein) lysate of A431 cells purchased from Upstate Biotechnologies (UK supplier = TCS, cat #12-110) or alternatively produced in-house from cells grown in tissue culture. The procedure of reacting a cellular or tissue lysate with a target protein is referred to as the Conditioning Reaction.

Lysates were produced by suspending approximately $5x10^6$ sub-confluent cells in 500ml of ice cold RIPA buffer (RIPAbuffer = 50mM Tris-HCL, pH7.4, 1%(v/v) NP-40, 0.25% sodium deoxycholate, 150mM NaCl, 1mM EGTA, 1mM PMSF, 1mM Na_3VO_4, 1mM NaF and 1ug/ml each of aprotinin, leupeptin and peptstatin). Cells had been previously washed of medium using ice cold PBS and were incubated with gentle shaking on ice for 15minutes in RIPA buffer. The lysate was cleared of insoluble material by centrifugation at 13,000 x g at 4°C for 10 minutes. The supernatant was removed and assayed for protein concentration using the bicinhoninic acid method and protocols provided by the supplier (Pierce, Chester UK, cat#23225). The lysate was divided into multiple aliquots, snap-frozen on dry ice and stored in liquid nitrogen before use in the conditioning reaction. Lysates from tissue samples and other cell types were processed in the same way.

The conditioning reaction was carried out at 37°C for 15 minutes and then placed on ice. The reaction was divided equally (by volume) into two tubes for detection of phosphorylated ERK-2 using either magnetic sorting (tube 1) or binding to microtitre plate wells (tube 2) in both cases using anti-histidine antibodies as capture reagent.

In tube 1, ERK-2 was captured using magnetic beads (Dynal, Wirral UK, cat#110.11) pre-coated with an anti-histidine antibody (Sigma, Poole UK, cat#H1029) using

coating conditions recommended by the supplier. The capture reaction was carried out at 4°C 1 hour with constant gentle mixing throughout to maintain the beads in suspension. Following capture, the beads were collected using a magnetic particle concentrator and washed extensively in PBS. Beads were resuspended in a final volume of 40ul and 10ul aliquots were taken into microtitre plate wells and individually reacted with dilutions of either anti-phosphorylated ERK antibody (Upstate Biotechnologies, cat#05-481), anti-phosphotyrosine-HRP conjugate monoclonal cocktail (Zymed, #136620), anti-phosphoserine /phosphothreoinine /phosphotyrosine polyconal preparation (Zymed, Cambridge UK, cat#90-0200) or a negative control antibody-HRP conjugate specific for human immunoglobulin light chains (The Binding Site, Birmingham UK, cat #APO015). Antibodies were incubated for 1 hour at 4°C before detection with respective secondary reagents/chromogenic substrates. Plates were read at 450nm.

In tube 2, ERK-2 was captured using microtitre plate wells which had been coated overnight with an anti-histidine antibody (Sigma, Poole UK, cat#H1029), and pre-blocked by 40 minute incubation with a solution PBS/5%(w/v) BSA at room temperature. The conditioning reaction was diluted to a total volume of 100ul using PBS and added to the microtitre plate. The capture reaction was for 1 hour at room temperature. Phosphorylated ERK-2 was detected in direct ELISA format using the panel of anti-bodies including a negative control reagent as given above.

Using both detection methods above, the presence of phosphorylated ERK-2 was unequivocally demonstrated following incubation with lysates from A431 cells.